GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 30, 2002, 15:55:16 ; Search time 76.19 Seconds
(without alignments)
949.857 Million cell updates/sec

Title: US-09-357-273A-2
Perfect score: 5139
Sequence: 1 MPARRLLLLLTLLLPGLGIF.....QPYYFHEPPEPQPPVTPDAL 977

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

al number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5093	99.1	977	22	AAM24098	Human EST encoded
2	2078.5	40.4	862	22	AAB65669	Novel protein kina
3	470	9.1	91	22	AAM32548	Peptide #6585 enco
4	452	8.8	86	22	AAM34296	Peptide #8333 enco
5	394	7.7	70	22	AAM19214	Peptide #5648 enco
6	394	7.7	70	22	AAM31867	Peptide #5904 enco
7	312	6.1	1108	20	AAY30046	Pancreatic eukaryo
8	294.5	5.7	971	19	AAW48896	Candida albicans C
9	294	5.7	718	18	AAW01537	Bovine ribonucleas
10	294	5.7	741	18	AAW12702	Human 2-5A-depende
11	294	5.7	741	20	AAW94679	Human 2-5A-depende
12	293	5.7	741	16	AAR82659	Human 2-5A-depende
13	293	5.7	741	16	AAR82660	Human 2-5A-depende
14	289	5.6	579	22	AAB48051	Signal transductio
15	288	5.6	741	15	AAR59076	2-5A-dependent RNA
16	280.5	5.5	647	19	AAW46576	Protein kinase LIM
17	279.5	5.4	626	21	AAB35805	Protein involved i
18	278.5	5.4	717	21	AAG46818	Arabidopsis thalia
19	276	5.4	520	18	AAW30918	Lily calcium/calmo
20	275	5.4	594	21	AAG30542	Arabidopsis thalia
21	275	5.4	649	21	AAG39133	Arabidopsis thalia
22	272.5	5.3	491	22	AAB48041	Signal transductio
23	268.5	5.2	499	22	AAU03512	Human protein kina
24	267.5	5.2	550	21	AAG33471	Arabidopsis thalia
25	267.5	5.2	601	21	AAG33470	Arabidopsis thalia
26	267.5	5.2	608	21	AAG33469	Arabidopsis thalia
27	266.5	5.2	586	21	AAG49798	Arabidopsis thalia
28	266.5	5.2	589	21	AAG49797	Arabidopsis thalia
29	266	5.2	679	15	AAR59077	2-5A-dependent RNA
30	266	5.2	679	16	AAR82661	Partial murine 2-5
31	266	5.2	679	18	AAW12703	Mouse 2-5A-depende
32	266	5.2	679	20	AAW94680	Mouse 2-5A-depende
33	265	5.2	510	21	AAG49366	Arabidopsis thalia
34	265	5.2	519	20	AAY24020	Mitogen-activated
35	265	5.2	522	21	AAG49365	Arabidopsis thalia
36	265	5.2	1068	20	AAY30048	Pancreatic eukaryo
37	265	5.2	1115	20	AAY30047	Pancreatic eukaryo
38	264.5	5.1	624	22	AAM41375	Human polypeptide
39	264.5	5.1	639	22	AAM39589	Human polypeptide
40	264.5	5.1	721	21	AAB18661	A human regulator
41	264.5	5.1	1230	19	AAW48895	Candida albicans C
42	264	5.1	978	22	AAB65604	Novel protein kina
43	263.5	5.1	681	20	AAY55940	Human PAK4 protein
44	263.5	5.1	681	21	AAB03967	Signal transductio
45	263.5	5.1	681	22	AAB20337	Human PAK5 full-le

ALIGNMENTS

RESULT 1
AAM24098
ID AAM24098 standard; Protein; 977 AA.
XX
AC AAM24098;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST encoded protein SEQ ID NO: 1623.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition.
XX
OS Homo sapiens.
XX
PN WO200154477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02687.
XX
PR 25-JAN-2000; 2000US-0491404.
PR 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.

```
DR   N-PSDB; AAH98757.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 20; Page 1099-1101; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a
CC   protein of the invention.
XX
SQ   Sequence   977 AA;


  Query Match             99.1%;  Score 5093;  DB 22;  Length 977;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  970;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 MPARRLLLLLTLLLPGLGIFGSTSTVTLPETLLFVSTLDGSLHAVSKRTGSIKWTLKEDP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mparrlllllltlllpglgifgststvtlpetllfvstldgslhavskrtgsikwtlkedp 60

Qy      61 VLQVPTHVEEPAFLPDPNDGSLYTLGSKNNEGLTKLPFTIPELVQASPCRSSDGILYMGK 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 vlqvpthveepaflpdpndgslytlgsknnegltklpftipelvqaspcrssdgilymgk 120

Qy     121 KQDIWYVIDLLTGEKQQTLSSAFADSLCPSTSLLYLGRTEYTITMYDTKTRELRWNATYF 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 kqdiwyvidlltgekqqtlssafadslcpstsllylgrteytitmydtktrelrwnatyf 180

Qy     181 DYAASLPEDEGDYKMSHFVSNGDGLVVTVDSESGDVLWIQNYASPVVAFYVWQREGLRKV 240
           |||||||||: ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 dyaaslpeddvdykmshfvsngdglvvtvdsesgdvlwiqnyaspvvafyvwqreglrkv 240

Qy     241 MHINVAVETLRYLTFMSGEVGRITKWKYPFPKETEAKSKLTPTLYVGKYSTSLYASPSMV 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 mhinvavetlryltfmsgevgritkwkypfpketeaksktptlyvgkystslyaspsmv 300

Qy     301 HEGVAVVPRGSTLPLLEGPQTDGVTIGDKGECVITPSTDVKFDPGLKSKNKLNYLRNYWL 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 hegvavvprgstlpllegpqtdgvtigdkgecvitpstdvkfdpglkskNklnylrnywl 360

Qy     361 LIGHHETPLSASTKMLERFPNNLPKHRENVIPADSEKKSFEEVINLVDQTSENAPTTVSR 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 lighhetplsastkmlerfpnnlpkhrenvipadsekksfeevinlvdqtsenapttvsr 420

Qy     421 DVEEKPAHAPARPEAPVDSMLKDMATIILSTFLLIGWVAFIITYPLSMHQQQQLQHQQFQ 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 dveekpahaparpeapvdsmlkdmatiilstflligwvafiityplsmhqqqqlqhqqfq 480

Qy     481 KELEKIQLLQQQQQQLPFHPPGDTAQDGELLDTSGPYSESSGTSSPSTSPRASNHSLCSG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 kelekiqllqqqqqqlpfhppgdtaqdgelldtsgpysessgtsspstsprasnhslcsg 540

Qy     541 SSASKAGSSPSLEQDDGDEETSVVIVGKISFCPKDVLGHGAEGTIVYRGMFDNRDVAVKR 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ssaskagsspsleqddgdeetsvvivgkisfcpkdvlghgaegtivyrgmfdnrdvavkr 600

Qy     601 ILPECFSFADREVQLLRESDEHPNVIRYFCTEKDRQFQYIAIELCAATLQEYVEQKDFAH 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ilpecfsfadrevqllresdehpnviryfctekdrqfqyiaielcaatlqeyveqkdfah 660

Qy     661 LGLEPITLLQQTTSGLAHLHSLNIVHRDLKPHNILISMPNAHGKIKAMISDFGLCKKLAV 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     661 lglepitllqqttsglahlhslnivhrdlkphnilismpnahgkikamisdfglwkklav 720

Qy     721 GRHSFSRRSGVPGTEGWIAPEMLSEDCKENPTYTVDIFSAGCVFYYVVSEGSHPFGKSLQ 780
           |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db     721 grhsfsrrsgvpgtegwiapemlsedckenptytvdifsagcvfyyvisegshpfgkslq 780

Qy     781 RQANILLGACSLDCLHPEKHEDVIARELIEKMIAMDPQKRPSANDVLKHPFFWSLEKQLQ 840
           |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db     781 rqanillgacsldclhpekhedviareliekmiamdpqkrpsakhvikhpffwslekqlq 840

Qy     841 FFQDVSDRIEKESLDGPIVKQLERGGRAVVKMDWRENITDPLQTDLRKFRTYKGGSVRDL 900
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     841 ffqdvsdriekesldgpivkqlerggravvkmdwrenitvplqtdlrkfrtykggsvrdl 900

Qy     901 LRAMRNKKHHYRELPAEVRETLGTLPDDFVCYFTSRFPHLLAHTYRAMELCSHERLFQPY 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 lramrnkkhhyrelpaevretlgtlpddfvcyftsrfphllahtyramelcsherlfqpy 960

Qy     961 YFHEPPEPQPPVTPDAL 977
           |||||||||||||||||
Db     961 yfheppepqppvtpdal 977


RESULT   2
AAB65669
ID   AAB65669 standard; Protein; 862 AA.
XX
AC   AAB65669;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Novel protein kinase, SEQ ID NO: 197.
XX
KW   Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW   immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW   dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW   immune disorder; cardiovascular disease; neurodegenerative disease;
KW   cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW   inflammatory pelvic disease; multiple sclerosis; psoriasis.
XX
OS   Homo sapiens.
XX
PN   WO200073469-A2.
XX
PD   07-DEC-2000.
XX
PF   26-MAY-2000; 2000WO-US14842.
XX
PR   28-MAY-1999;   99US-0136503.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Martinez R,  Whyte D,  Sudersanam S;
XX
DR   WPI; 2001-032161/04.
DR   N-PSDB; AAF44697.
XX
PT   Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT   treating immune-related diseases and disorders, cardiovascular disease,
PT   neurodegenerative diseases and/or cancers -
XX
PS   Claim 10; Fig 1; 310pp; English.
XX
CC   The present sequence is a novel protein kinase. The novel protein kinases
CC   and the nucleic acids that encode them may be used in the treatment and
CC   diagnosis of diseases associated with inappropriate kinase expression
CC   such as immune-related diseases and disorders, cardiovascular disease,
CC   neurodegenerative diseases and/or cancers. The nucleic acids and
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays. The kinase polypeptides may be used as antigens in the production
CC   of antibodies of kinase expression and activity. Anti-kinase antibodies
CC   and kinase antagonists may also be used to down regulate kinase
CC   expression and activity. Diseases related to kinase expression and
CC   activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC   disorders, complications of organ transplantation, myocardial infarction,
CC   immune disorders, cardiomyopathies, strokes, renal failure,
CC   oxidative-stress related disorders, chronic inflammatory bowel disease,
```

```
CC   chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC   osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC   reproductive disorders.
XX
SQ   Sequence   862 AA;


  Query Match             40.4%;  Score 2078.5;  DB 22;  Length 862;
  Best Local Similarity   46.8%;  Pred. No. 4.8e-159;
  Matches  464;  Conservative  113;  Mismatches  254;  Indels  161;  Gaps   18;

Qy      2 PARRLLLLLTLLLPGLGIFGSTSTVTLPETLLFVSTLDGSLHAVSKRTGSIKWTLKEDPV 61
          |  || | |      ||          || || ||||||||||:||:|| :||||::|||
Db     10 pwprlglqlqfaalllgtlspqvhtlrpenlllvstldgslhalskqtgdlkwtlrddpv 69

Qy     62 LQVPTHVEEPAFLPDPNDGSLYTLGSKNNEGLTKLPFTIPELVQASPCRSSDGILYMGKK 121
          :: | :| | ||| || ||||| ||::  :|| |||||||||| |||||||||: | |:|
Db     70 iegpmyvtemaflsdpadgslyilgtqkqqglmklpftipelvhaspcrssdgvfytgrk 129

Qy    122 QDIWYVIDLLTGEKQQTLSSAFADSLCPSTSLLYLGRTEYTITMYDTKTRELRWNATYFD 181
          || |:|:|  :|| | ||::       |||  ||:|||:||:||:| :   ||||
Db    130 qdawfvvdpesgetqmtltte-----gpstprlyigrtqytvtmhdprapalrwnt---- 180

Qy    182 YAASLPEDEGDYKMSHFVSNGDGLVVTVDSESGDVLWIQNYASPVVAFYVWQREGLRKVM 241
                                                                  |::
Db    181 ------------------------------------------------------rqlp 184

Qy    242 HINVAVETLRYLTFMSGEVGRITKWKYPF--PKETEA-----KSKLTPTLYVGKYSTSLY 294
          |: :| :|| :|    | :      : |   |::|        ::|  |||||  |  |
Db    185 hltlardtlhflalrwghi------rlpasgprdtatlfstldtqllmtlyvgkdetgfy 238

Qy    295 ASPSMVHEGVAVVPRGSTLPLLEGPQTDGVTIGDKGECVITPSTDVKFDPGLKSKNKLNY 354
           | ::|| |||:|||| ||   :|| || ||:   ||   :||| |::  |  :
Db    239 vskalvhtgvalvprgltlapadgpttdevtlqvsgeregspstavrypsgsva------ 292

Qy    355 LRNYWLLIGHHETPLSASTKMLERFPNNLPKHRENVIPADSEKKS-FEEVINLVDQTSEN 413
          | : |||||||| |    | ||   |       |   | :::  : | |:::|
Db    293 lpsqwlligh helppvlhttmlrvhptlgsgtaetrppentqapafflellsl------- 345

Qy    414 APTTVSRDVEEKPAHAPARPEAPV-DSML----KDMATIILSTFLLIGWVAFIITYPLSM 468
              ||   ||   :   ||   || |    :|: .  |:  || ||: |:
Db    346 -----sr---eklwdselhpeektpdsylglgpqdllaasltavllggwilfv------- 390

Qy    469 HQQQQLQHQQFQKELEKIQLLQQQQQQLPFHPP--GDTAQDGELLDTSGPYSESSGTSSP 526
                       :: : ::||: |  |      :|| : |          || |
Db    391 -----------------mrqvvekqqetplapadfahisqdaqsl--------hsgas-- 423

      527 STSPRASNHSLCSGSSASKAGSSPSLEQDDGDEETSVVIVGKISFCPKDVLGHGAEGTIV 586
          | |   | | |  ::    |  ||        : :|||||| |||||| || || |
      424 ----rrsqkrlqspskqaqpldd peaeq--------ltvvgkisfnpkdvlgrgaggtfv 471

Qy    587 YRGMFDNRDVAVKRILPECFSFADREVQLLRESDEHPNVIRYFCTEKDRQFQYIAIELCA 646
          :|| |: | ||||||:| |||    ||||||:||| |||||||||||:  || |||:|||
Db    472 frgqfegravavkrllrecfglvrrevqllqesdrhpnvlryfctergpqfhyialelcr 531

Qy    647 ATLQEYVEQKDFAHLGLEPITLLQQTTSGLAHLHSLNIVHRDLKPHNILISMPNAHGKIK 706
          |:|||||  |  . ||||  :|||  |||||||||:|||||||| |||| |:: |  :
Db    532 aslqeyvenpdldrgglepevvlqqlmsglahlhslhivhrdlkpgnilitgpdsqglgr 591

Qy    707 AMISDFGLCKKLAVGRHSFSRRSGVPGTEGWIAPEMLSEDCKENPTYTVDIFSAGCVFYY 766
           ::|||||||||  || |||  ||:||||||:|||:|     ::||  ||||||||||||
Db    592 vvlsdfglckklpagrcsfslhsgipgtegwmapellqllppdsptsavdifsagcvfyy 651

Qy    767 VVSEGSHPFGKSLQRQANILLGACSLDCLHPEKHEDVIARELIEKMIAMDPQKRPSANDV 826
          |:| |||||| || |||||| ||   |  |  | |: |:||:|:  |::  || |||| |
Db    652 vlsggshpfgdslyrqaniltgapclahleeevhdkvvardlvgamlsplpqprpsapqv 711

Qy    827 LKHPFFWSLEKQLQFFQDVSDRIEKESLDGPIVKQLERGGRAVVKMDWRENITDPLQTDL 886
          | |||||  ||||||||||| :||||   |:|: || || |||: :| |:|: ||||||
Db    712 lahpffwsrakqlqffqdvsdwlekeseqeplvraleaggcavvrdnwhehismplqtdl 771

Qy    887 RKFRTYKGGSVRDLLRAMRNKKHHYRELPAEVRETLGTLPDDFVCYFTSRFPHLLAHTYR 946
          ||||:||| |||||||:|||||||||| |||: || :|| || |||:||| || ||:|
Db    772 rkfrsykgtsvrdllravrnkkhhyrelpvevrqalgqvpdgfvqyftnrfprlllhthr 831

Qy    947 AMELCSHERLFQPYYFHEPPE-----PQPPVT 973
              ||  |: | || |||   ||:     | |  |
Db    832 amrscaseslflpyy---ppdseairpcpgat 860


RESULT   3
AAM32548
ID   AAM32548 standard; Protein; 91 AA.
XX
AC   AAM32548;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Peptide #6585 encoded by probe for measuring placental gene expression.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 27; SEQ ID No 32817; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP:
CC   see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC   such probe. The probes are useful for producing a microarray for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from human placenta. The probes are useful for antenatal diagnosis of
CC   human genetic disorders.
XX
SQ   Sequence   91 AA;


  Query Match              9.1%;  Score 470;  DB 22;  Length 91;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-30;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    467 SMHQQQQLQHQQFQKELEKIQLLQQQQQQLPFHPPGDTAQDGELLDTSGPYSESSGTSSP 526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 smhqqqqlqhqqfqkelekiqllqqqqqqlpfhppgdtaqdgelldtsgpysessgtssp 60

Qy    527 STSPRASNHSLCSGSSASKAGSSPSLEQDDG 557
          |||||||||||||||||||||||||||||||
Db     61 stsprasnhslcsgssaskagsspsleqddg 91


RESULT   4
AAM34296
ID   AAM34296 standard; Protein; 86 AA.
```

```
XX
AC   AAM34296;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Peptide #8333 encoded by probe for measuring placental gene expression.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
     26-MAY-2000; 2000US-0207456.
     30-JUN-2000; 2000US-0608408.
     03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 27; SEQ ID No 34565; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP:
CC   see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC   such probe. The probes are useful for producing a microarray for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from human placenta. The probes are useful for antenatal diagnosis of
CC   human genetic disorders.
XX
SQ   Sequence   86 AA;


 Query Match           8.8%;  Score 452;  DB 22;  Length 86;
 Best Local Similarity   100.0%;  Pred. No. 4.5e-29;
 Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     195 MSHFVSNGDGLVVTVDSESGDVLWIQNYASPVVAFYVWQREGLRKVMHINVAVETLRYLT 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshfvsngdglvvtvdsesgdvlwiqnyaspvvafyvwqreglrkvmhinvavetlrylt 60

Qy     255 FMSGEVGRITKWKYPFPKETEAKSKL 280
           |||||||||||||||||||||||||
Db      61 fmsgevgritkwkypfpketeakskl 86


RESULT   5
AAM19214
ID   AAM19214 standard; Protein; 70 AA.
XX
AC   AAM19214;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Peptide #5648 encoded by probe for measuring cervical gene expression.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 27; SEQ ID No 24040; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC   by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC   can be used to produce a single exon microarray, which can be used for
CC   measuring human gene expression in a sample derived from human cervical
CC   epithelial cells. By measuring gene expression, the probes are therefore
CC   useful in grading and/or staging of diseases of the cervix, notably
CC   cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   70 AA;


 Query Match           7.7%;  Score 394;  DB 22;  Length 70;
 Best Local Similarity   98.6%;  Pred. No. 1.6e-24;
 Matches   69;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     908 KHHYRELPAEVRETLGTLPDDFVCYFTSRFPHLLAHTYRAMELCSHERLFQPYYFHEPPE 967
           ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db       1 khhyrelpaevretlgslpddfvcyftsrfphllahtyramelcsherlfqpyyfheppe 60

Qy     968 PQPPVTPDAL 977
           ||||||||||
Db      61 pqppvtpdal 70


RESULT   6
AAM31867
ID   AAM31867 standard; Protein; 70 AA.
XX
AC   AAM31867;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Peptide #5904 encoded by probe for measuring placental gene expression.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
```

```
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 27; SEQ ID No 32136; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP:
CC   see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC   such probe. The probes are useful for producing a microarray for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from human placenta. The probes are useful for antenatal diagnosis of
CC   human genetic disorders.
XX
SQ   Sequence   70 AA;


  Query Match             7.7%;  Score 394;  DB 22;  Length 70;
  Best Local Similarity   98.6%;  Pred. No. 1.6e-24;
  Matches   69;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    908 KHHYRELPAEVRETLGTLPDDFVCYFTSRFPHLLAHTYRAMELCSHERLFQPYYFHEPPE 967
          |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db      1 khhyrelpaevretlgslpddfvcyftsrfphllahtyramelcsherlfqpyyfheppe 60

Qy    968 PQPPVTPDAL 977
          ||||||||||
Db     61 pqppvtpdal 70


RESULT   7
AAY30046
     AAY30046 standard; Protein; 1108 AA.

     AAY30046;
XX
DT   04-OCT-1999  (first entry)
XX
DE   Pancreatic eukaryotic translation initiation factor-2 alpha kinase.
XX
KW   Pancreatic eukaryotic translation initiation factor-2 alpha kinase; PEK;
KW   eukaryotic translation initiation factor-2 alpha; pancreatic islet;
KW   drug discovery; drug development.
XX
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   20
FT                   /note= "consensus N-myristylation site"
FT   Modified-site   44
FT                   /note= "consensus N-myristylation site"
FT   Region          517..532
FT                   /note= "hydrophobic region; potential transmembrane
FT                          region"
XX
PN   WO9938994-A1.
XX
PD   05-AUG-1999.
XX
PF   12-JAN-1999;   99WO-US00623.
XX
PR   25-NOV-1998;   98US-0109992.
PR   29-JAN-1998;   98US-0073031.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Shi Y;
XX
DR   WPI; 1999-469338/39.
DR   N-PSDB; AAX86563.
XX
PT   Novel rat and human pancreatic eukaryotic translation initiation
PT   factor 2alpha kinase useful for drug discovery and development
XX
PS   Claim 7; Page 46-50; 77pp; English.
XX
CC   The present sequence represents a pancreatic eukaryotic translation
CC   initiation factor-2 alpha kinase (PEK). PEK phosphorylates eukaryotic
CC   translation initiation factor-2 alpha. PEK polynucleotides was
CC   cloned from pancreatic islet DNA libraries. The PEK nucleic acids
CC   and protein can be used as tools for drug discovery and development.
XX
SQ   Sequence   1108 AA;


  Query Match             6.1%;  Score 312;  DB 20;  Length 1108;
  Best Local Similarity   18.8%;  Pred. No. 7.2e-16;
  Matches  212;  Conservative  186;  Mismatches  379;  Indels  348;  Gaps   51;

Qy      2 PARRLLLLLTLLLP----------------------GLGIFGS-TSTVTLP-------- 29
          |  | |||| |||                          |||   : ||   :|
Db      7 prprallllflllgcaagisavararsllaptsdtafglgaaaaptsaarvpavataevt 66

Qy     30 ------------------------------ETLLFVSTLDGSLHAV-SKRTGSIKWTLK 57
                                          :|: :||||| : |: ::  |  :| |
Db     67 vedaealpaasgeqesratesdddvelrprgrslviistldgriaaldaenhgkkqwdld 126

Qy     58 EDPVLQVPTHVEEPAFLPD----PN-DGSLYTLGSKNNEGLTKLPFTIPELVQASPCRSS 112
              | : : :|    :    |: || |:    :: | :  :|||:  |:::|  :
Db    127 vgsgslvssslskpevfgnkmiipsldgdlfq-wdrdresmeavpftveslless-ykfg 184

Qy    113 DGILYMGKKQDIWYVIDLLTGEKQQTLSSAF------ADSLCPSTSLLYLGRTEYTITMY 166
          | :: :| |    | :   :| | : : ||       :| :     :| | ||: |:
Db    185 ddvvlvggksltty glsaysg-klryicsalgcrrwdsdemeeeedilllqrtqktvrav 243

Qy    167 DTKTRELRWN----------------ATYFDYAASLPEDEGDYKMSH----------F 198
          |  ::   :||               | : :    |  :: | |:
Db    244 gprsgsekwnfsvghfelryipdmetragfiestfklggnkedskiisdveeqdvdtvik 303

Qy    199 VSNGDGLVVTVDSESGDVLWIQNYASPVVAFYVWQREGLRKVMHINVAVETLRYLTFMSG 258
          ||  |  |:    : | : |   : :|: : ::  |:|  ||: |::  :|
Db    304 vsvadwkvmafskkggrleweyqfctpiasawl-vrdg--kvipislfddt--------- 351

Qy    259 EVGRITKWKYPFPKET---------EAKSKLTPTLYVGKYSTSLYASPSM---------- 299
          |      |   :|          |:     ::|:| |   ||   |:
Db    352 |-------sytaneevledeediveaargatensvylgmyrgqlylqssvrvsekfptrp 403

Qy    300 |-----VHEGVAVVPRGST--LPLLEGPQTDGVTIG--DKGECVITPS-TDVKFDPGLKSK 349
          |     |:   |::|  :    ||:  |    | :|  :  :|:     :  ::  |  |
Db    404 kalesvngesaiiplptikwkplihspsrtpvlvgsdefdkclsndkysheeysngalsi 463

Qy    350 NKLNYLRNYWLLIGHHETPLSASTKMLERFPNNLPKHRENVIPADS------EKKSFEEV 403
          |:  |   |:|     |     ||::  || :: | : :|:   |        |:| :
Db    464 lqypydngyylpyykrern-krstqitvrflds-physknirkkdpilllhwwkeifgti 521

Qy    404 INLVDQTSENAPTTVSRDV-------EEKPAHAPARPEAPVDSMLKDMATIILSTFLLIG 456
          | :  |     | : | :       : | :    : |:  ||:  | :    :    |
Db    522 llcivat-----tfivrrlfhpqphrqrkesetqcqteskydsvsadnsdnswndikhsg 576

Qy    457 WVAFIIT--------------------------YPL----------------SMHQQQ 472
          :|:  :|                          | :                :
```

```
Db    577 yvsryltdfepiqcmgrggfgvvfeaknkvddcnyaikrirlpnrelarekvmrevkala 636

Qy    473 QLQH--------------QQFQKELEKIQLLQQQQQQLPFHPPGDTAQDGELLDTSGPY 517
          :|:|               :::|:|:::|  |: :    |  |         :    |:
Db    637 klehpgivryfnawletppekwqeemdei-wlkdestdwplsspspmdapsvkirqmdpf 695

Qy    518 SESSGTSSPSTSP-RASNHSL---CSGSSASKAGSSPSLE------------------Q 554
          |       : || |: : |:   |  :|:|::  || ||                   |
Db    696 stkeqieviapspersrsfsvgiscgrtsssesqfsp-lefsgtdcgdnsdsedaahnlq 754

Qy    555 D------------DGDEETSVVIVGKISFCPKDVLGH-----GAEGTIVYRGM-FDN-- 593
          |            |||:|          || :   :     |   :||:     ||
Db    755 dscltdcdmedgtvdgddegh-----sfelcpseaspytrsregtsssivfedsgcdnas 809

Qy    594 --RDVAVKRI---------LPECFSFADREVQ--LLRESDEHP---------NVIRYFCT 631
            |  : |:          | |    : |      |  |   |          | :
Db    810 skedprmnrlhnghhyvnkltefkhsssrssseatlstsptrpttlsldftrntvdrlqp 869

Qy    632 EKDRQFQYIAIELC-AATLQEYVEQK----DFAHLGLEPITLLQQTTSGLAHLHSLNIVH 686
            : : || ::||    |:::: ::    |  |      | |  |    :  |||  ::|
Db    870 sspkvylyiqmqlcrkenlkdwmnrrcsmedrehrvclhifl--qiaeavqflhskglmh 927

Qy    687 RDLKPHNILISMPNAHGKIKAMISDFGLC---------KKLAVGRHSFSRRSGVPGTEGW 737
          ||||| ||  :| :  .    : ||||          : :     :::  :|  ||: :
Db    928 rdlkpsnifftmddv----vkvgdfglvtamdqdeeeqtvltpmpayathtgqvgtkly 982

Qy    738 IAPEMLSEDCKENPTYTVDIFSAGCVFYYVVSEGSHPFGKSLQRQANILLGACSLDC--- 794
          ::|| :  :   | :: |||||  | : : ::    :||   ::|    |    :|
Db    983 mspeqihgn---nyshkvdifslglilfell----ypfstqmer-vrtltdvrnlkfppl 1034

Qy    795 ---LHPEKHEDVIARELIEKMIAMDPQKRPSANDVLKHPFFWSLE 836
              :|::|       ::: |::  | :|| | |::::  | :||
Db   1035 ftqkypqeh------mmvqdmlspspmerpeatdiienavfenle 1073


RESULT   8
AAW48896
ID   AAW48896 standard; Protein; 971 AA.
XX
AC   AAW48896;
XX
DT   13-OCT-1998  (first entry)
XX
DE   Candida albicans CaCLA4 protein.
XX
KW   CaCLA4; protein kinase; Ste20p family; screening; virulence;
KW   hyphal formation; pathogenic fungi; inhibitor; inflammation;
KW   antimycotic.
XX
OS   Candida albicans.
XX
PN   WO9818927-A1.
XX
PD   07-MAY-1998.
XX
PF   29-OCT-1997;   97WO-CA00809.
XX
PR   30-OCT-1996;   96US-0029458.
XX
PA   (CANA ) NAT RES COUNCIL CANADA.
XX
PI   Leberer E,  Thomas DY;
XX
DR   WPI; 1998-272222/24.
DR   N-PSDB; AAV32554.
XX
PT   In vitro screening test for agents that inhibit Candida genes
PT   involved in virulence - and transition to hyphal form, potentially
PT   useful as antimycotic agents
XX
PS   Disclosure; Fig 7; 79pp; English.
XX
CC   The sequence is that of the CaCLA4 protein which can be used
CC   in the development of an in vitro screening test for compounds
CC   that inhibit biological activity of the protein and a system for
CC   measuring its activity. The protein is involved in virulence and
CC   hyphal formation. Inhibitors are potentially useful for rendering
CC   pathogenic fungi (any species in which hyphal induction by kinase
CC   occurs) avirulent and/or to treat inflammation.
XX
SQ   Sequence   971 AA;


  Query Match             5.7%;  Score 294.5;  DB 19;  Length 971;
  Best Local Similarity   19.7%;  Pred. No. 1.5e-14;
  Matches  186;  Conservative  129;  Mismatches  279;  Indels  349;  Gaps   40;

Qy    174 RWNATYF---DYAASLPEDEGDYKMSHFVSNGDGLVVTVDSESGDV---LWIQN------ 221
          |||  :    |  :       ||   : |:|:      :| : |:   |:: |
Db     86 rwnkrfmvindktlnf------ykqepyssdgns-----nsntpdlsfplylinninlkp 134

Qy    222 ---YASPVVAFYVWQREGLRKVMHINVAVETLR-YLTFMSGEVGRITKWKYPFPKETEAK 277
             |:    :| :  :   : ::   ::|:|   || ::                   :|
Db    135 nsgysktsqsfeivpknnnksil---isvktnndyldwl-----------------daf 173

Qy    278 SKLTPTLYVGKYSTSLYASPSMVHEGVAVVPRGS---------TLPLLEGPQTDGVTIGD 328
          :   | : :|: :: :  | |  |  :  :  ||         |  ||      |   |
Db    174 ttkcplvqigennsgv--ssshphlqiqhltngslngnsssssptsglllsssvltggnsgv 231

Qy    329 KGECVITPSTDVKFDPGLKSKNKLNYLRNYWLLIGHHE---------------------- 366
          |    |   | |||    :       | : |  :  |
Db    232 sgpinfthkvhvgfdp---asgnftglpdtwkslqhskitnedwkkdpvavievlefys 288

Qy    367 ----------TP-----LSASTKMLERFPNNLPKHRENVIPADSEKKSFEEVINLVDQTS 411
                     ||     ::: |      |||    : ||  |           || :
Db    289 dinggnsaagtpigspminsktnnnnndpnnysstknnvqea----------nlqewvk 337

Qy    412 ENAPTTVSRDVEEKPAHAPARPEAPV------------------------------DS 439
            | :|||   : ||: |  :|  |                              ::
Db    338 ppakstvs---qfkpsraapkpptpyhltqlngsshqhtsssgslpssgnnnnnnstnnn 394

Qy    440 MLKDMATI--ILSTFLLI----------GWVAFIITYPLSMHQ---------------- 470
            |::: :  :::   ||          |  :   ||    ||
Db    395 ntknvsplnnlmnkseliparrappppptsgtss--dtysnknhqdrsgyeqqrqqrtdss 452

Qy    471 QQQLQHQQFQKELEKIQLLQQQQQQLPFHPPGDT---AQDGELLDTSGPYSE-SSGTSSP 526
          ||| | :| | : :  |  || || |  |  | :    |    | :||| :: :|| | |
Db    453 qqqqqqkqhqyqqksqqqqqqpqqplslhqggtshipkqvpptlpssgpptqaasgksmp 512

Qy    527 S-------TSPRASNHSLCSGSS-------------------------ASKAGSSPSLE 553
          |            :|:   ||:                          |||  | || :
Db    513 skihpdlkiqqgtnnyikssgtdanqvdgdakqfikpfnlqlkksqqqlaskqpsppssq 572

Qy    554 QDDGDEETS---------------------------------------------------- 562
          |      ||
Db    573 qqqqkpmtshglmgtshsvtkplnpvndpikplnlkssksksealnetlgvlktpsptdks 632

Qy    563 ---------------------------VVIVGKISFC-----PKDVL------GHGAEG 583
                                     : |: |:        || :       | || |
Db    633 nkptapasgpavtktakqlkkererlndlqiiaklktvvnnqdpkplfrivekagqgasg 692

Qy    584 TIVYRGMF--DNRDVAVKRILPECFSFADREVQ----------LLRESDEHPNVIRYFCT 631
            :    |   :|| :|:|::        | : |          |: :  :| |:: :  :
Db    693 nvylaemikdnnrkiaikqm--------dldaqprkeliineilvmkdsqhknivnflds 744

Qy    632 E--KDRQFQYIAIELCAATLQEYVEQKDFAHLGLEPITLLQQTTSGLAHLHSLNIVHRDL 689
            | :   |   :   :| | :|  ||     :  |:  :|  || ||| .:|:|||:
Db    745 yligdnelwvimeymqggslteiienndfklnekqiaticfetlkglqhlhkkhiihrdi 804

Qy    690 KPHNILISMPNAHGKIKAMISDFGLCKKLAVGRHSFSRRSGVPGTEGWIAPEMLSEDCKE 749
          |  |:|:   :|:| :|  |:||| | ||   |   ::|: : ||  |:|||:: :  ||
Db    805 ksdnvll---daygnvk--itdfgfcakltdqr---nkratmvgtpywmapevvkq--ke 854

Qy    750 NPTYTVDIFSAGCVFYYVVSEGSHPFGKSLQRQANILLGACSLDCL-HPEKHEDVIAREL 808
              ||::| | :   :: ||  |:      :|  |:       |  ||   : | ::
Db    855 yde-kvdvwslgimtiemi-egeppylneeplkalyliatngtpklkkpellsnsi-kkf 911
```